Mapping U.S. FDA National Drug Codes to Anatomical-Therapeutic-Chemical Classes using RxNorm

Fabrício S. P. Kury, MD[1], Olivier Bodenreider, MD, PhD[1]
[1]U.S. National Library of Medicine, U.S. National Institutes of Health, Bethesda, MD, USA

Background

The U.S. Food and Drug Administration National Drug Codes (NDCs) are the official identifiers for drug products in the U.S. and therefore commonly appear in large drug prescription datasets from this country. To analyze these datasets from a clinical perspective, it is often useful to abstract away from details such as manufacturer, dosage form and packaging information, and to group prescriptions by drug classes. The Anatomical-Therapeutic-Chemical[1] drug classification system provides a 5-level aggregation scheme for drugs. However, neither WHO nor FDA provides a map from NDC to ATC. Here we demonstrate how to use the U.S. National Library of Medicine RxNorm[2] application programming interface (API) to automatically create such a map. We apply this by mapping 71,309 NDCs found in 1.015 billion Medicare Part D claims from 2006 to 2013, and 134,580 NDCs found in a commercial all-payer claims dataset from Partners Healthcare© from 2011 to 2012.

Methods and Results

We produced an R script and made it freely available under a Creative Commons *Attribution-NonCommercial-ShareAlike 4.0 International* license in the GitHub repository at https://github.com/fabkury/ndc_map. The script uses caching and parallel processing to greatly improve efficiency. The script ingests a list of NDCs, queries the RxNorm API to obtain their RxNorm Concept Unique Identifiers (RxCUIs), then queries the API again to obtain the fourth-level ATC class or classes (ATC-4) of each RxCUI. Therefore, our mapping can fail at two stages: when NDCs are not recognized by RxNorm, and when an RxNorm drug has no associated ATC-4 class. Tables 1 and 2 provide statistics about the mapping of the two datasets under investigation.

Table 1. Results of the mapping process.

	Medicare	All-payer
Mapped prescriptions	97.6%	*Not available.*
Mapped NDCs	55,565 (77.9%)	97,961 (72.8%)
NDCs with no RxCUI	10,413 (14.6%)	29,160 (21.7%)
RxCUI but no ATC-4	5,331 (7.5%)	7,459 (5.5%)
Unique NDC—ATC-4	114,069	187,426
Unique ATC-4 classes	517 *(58.62%)*	540 *(61.2%)*

Table 2. Tallies of ATC-4 classes per NDC.

ATC-4 per NDC	Medicare	All-payer
0	15,744 (22.1%)	36,619 (27.2%)
1 to 2	45,266 (63.5%)	78,414 (58.3%)
3 to 8	8,887 (12.5%)	17,667 (13.1%)
9 to 21	1,412 (2%)	1,880 (1.4%)
≥ 22	0	0

Discussion and Conclusion

Medicare claims contained a smaller set of NDCs which were proportionally better covered by RxNorm. In the Medicare dataset, only 77.9% of the NDCs were mapped to at least one ATC class, but those NDCs represented 97.6% of all claims. From an informal analysis of the Medicare data, the majority of the non-mapped claims were for over-the-counter medications or non-drug supplies (e.g. syringes, gauze). In both datasets, individual NDCs were ambiguously associated with up to 21 ATC-4 classes for two reasons: some drugs have more than one active ingredient with different ATC classes; and ATC can provide multiple classes for the same ingredient depending on its usage. For example, *miconazole* has 6 ATC level 4 classes across oral use, gynecological use, otological use, and so on. If not addressed, the NDC-to-ATC ambiguity can create issues in analyzing prescription datasets. **Conclusion:** Our approach leveraging RxNorm provides a simple and scalable solution for mapping NDCs to ATC classes. The ambiguity of some mappings is the biggest concern for researchers using the NDC-to-ATC map.

References

1. World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2016. 2016. Accessed from http://www.whocc.no/atc_ddd_index/ on March 9, 2017.
2. U.S. National Library of Medicine. RxNorm APIs. Accessed from https://rxnav.nlm.nih.gov/APIsOverview.html on March 9, 2017.